

УДК 655.244.07

## **Дискриминантный анализ как метод прогнозирования удобочитаемости шрифтов**

Токарь О. В

*Белорусский государственный технологический университет, факультет  
принттехнологий и медиакоммуникаций, Свердлова, 13а-117-4, Минск,  
Беларусь, 220050*

**Аннотация.** Статья посвящена разработке функций для прогнозирования удобочитаемости шрифта, влияющей на качество восприятия материала в печатном и электронном виде, методом дискриминантного анализа на основе геометрических параметров шрифта.

**Ключевые слова:** шрифт, удобочитаемость, дискриминантный анализ.

## **Discriminant analysis as a method of predicting the readability of fonts**

Tokar O. V.

*Belarusian state technological University, Faculty of printtechnology and media  
communications, Sverdlov, 13A-117-4, Minsk, BELARUS, 220050*

**Abstract.** The article is devoted to developing features to predict the readability of the font influencing the quality perception of the material in printed and electronic form, the method of discriminant analysis based on geometric parameters of the font.

**Keywords:** the font, readability, discriminant analysis.

### **Введение**

Известно, что удобочитаемость шрифтов и качество шрифтового оформления печатных и электронных источников информации связаны с качеством ее вос-

приятия и усвоения реципиентом. В связи с этим актуальна разработка новых методик для установления удобочитаемости материала и параметров, позволяющих ее регулировать.

Цель работы заключается в разработке классификационных функций для прогнозирования удобочитаемости шрифтов, измеренной методиками скорости чтения и парных сравнений.

## 1. Метод

Для достижения цели может быть использован дискриминантный анализ, включающий несколько тесно связанных статистических процедур, которые можно разделить на методы интерпретации межгрупповых различий — дискриминации и методы классификации наблюдений по группам [1].

Были отобраны следующие 25 шрифтов, разработанные фирмой «ПараТайп»: AdonisCTT; AvantGardeGothicCTT; BookmanCTT; BellGothicCTT BT; CharterCTT; ErasLightCTT; FuturaFuturisCTT; GeoSlb712CTT Md BT; KabelCTT Medium; KisSCCTT BT; NewBaskervilleCTT; NewtonCTT; OCRF-RegularCTT; OctavaCTT; PragmaticaCTT; RaleighCTT BT; CooperCTT Lt BT; FranklinGothicBookCTT; TextBookCTT; Gothic725CTT Bd BT; Humanist531CTT BT; SwiftCTT; ZapfElliptical711CTT BT; OfficinaSansMediumCTT; OriginalGaramondSCCTT BT., а также шрифт Times New Roman из-за привычки к этому шрифту большинства читателей.

Для каждого шрифта рассчитаны геометрические параметры: отношение кегля шрифта к высоте знака, пропорциональность, контрастность и т. д. Методики оценки удобочитаемости и описание геометрии шрифтов приведены в [2]. Влияние геометрических параметров шрифтов на степень их удобочитаемости, измеренной различными методиками, отражено в публикациях [3, 4, 5].

## 2. Результаты

Исходная информация представляет собой данные о шрифтах, относящихся к группе с высокой удобочитаемостью и с низкой по двум методикам.

В группу с высокой удобочитаемостью по методу парных сравнений включены 11 шрифтов: AdonisCTT, BookmanCTT, CharterCTT, NewtonCTT, OctavaCTT, PragmaticaCTT, Times New Roman, CooperCTT Lt BT, FranklinGothicBookCTT, SwiftCTT, ZapfElliptical711CTT BT.

Группу с низкой удобочитаемостью по методу парных сравнений образовали 15 шрифтов: AvantGardeGothicCTT, BellGothicCTT BT, ErasLightCTT, FuturaFuturisCTT, GeoSlb712CTT Md BT, KabelCTTMedium, KisSCCTT BT, NewBaskervilleCTT, OCRF-RegularCTT, RaleighCTT BT, Gothic725CTT Bd BT, Humanist531CTT BT, TextBookCTT, OfficinaSansMediumCTT, OriginalGaramondSCCTT BT.

Поскольку число дискриминантных функций меньше числа групп на единицу, в данном случае при помощи пакета Statgraphics была получена одна дискриминантная функция со следующими характеристиками: каноническая корреляция (0,89) и оценки уровня значимости функции по критериям Лямбда Уилкса (0,21) и хи-квадрат (27,0).

Значение статистики Уилкса в интервале от 0 до 1 свидетельствует о неплохой дискриминации, т. е. о том, что задействованные переменные эффективно участвуют в различении групп. Большая величина канонической корреляции соответствует тесной связи дискриминантной функции с группами. Хи-квадрат имеет большое значения, что говорит о том, что она обладает хорошей дискриминативной способностью.

Интерпретации различия между группами шрифтов с высокой и невысокой удобочитаемостью способствуют коэффициенты дискриминантной функции (таблица 1).

Таблица 1. Коэффициенты дискриминантной функции при классификации шрифтов по методу парных сравнений

Параметр	Коэффициент	Параметр	Коэффициент
Пропорциональность «н»	10,18	Контрастность «а»	-2,72
Контрастность «н»	1,85	Max/Min «А»	-0,73
Засечки «н»	-3,55	Кегль/Высота «а»	-16,22
Площадь «н»	-0,03	Штрих/Просвет «а»	14,48
Периметр «н»	1,15	Площадь «а»	-0,0003
Кегль/Высота «н»	18,28	Периметр «а»	-0,09
Штрих/Просвет «н»	-3,77	Константа	-122,89
Пропорциональность «а»	-40,02		

Наибольший вклад в дискриминацию в данном случае вносят параметры пропорциональности и отношения кегля к высоте для знаков «н» и «а», а также отношение основного штриха к внутрибуквенному просвету знака «а», т. е. эти параметры влияют на удобочитаемость шрифта.

Точность классификации шрифтов достаточно высокая (таблица 2). Из 26 шрифтов в группе с высокой удобочитаемостью по методу парных сравнений изначально находятся 11 объектов, в группе шрифтов с низкой удобочитаемостью состоит 15 объектов. Точность распознавания объектов первой группы составила 100 % (распознаны все 11 шрифтов), второй — 93 % (из 15 шрифтов распознаны 14).

Таблица 2. Классификационная таблица (метод парных сравнений)

Исходная группа	Предсказанная группа		Итого
	1	2	
1	11 (100 %)	0	11
2	1 (7 %)	14 (93 %)	15

Для классификации новых объектов, т. е. прогнозирования удобочитаемости (по методу парных сравнений) новых шрифтов только лишь по геометрическим параметрам были получены следующие классификационные функции:

$$F1 = -143389,0 + 3532,27 \cdot X1 - 570,90 \cdot X2 - 822,60 \cdot X3 - 22,72 \cdot X4 + 926,64 \cdot X5 + 21773,7 \cdot X6 + 1103,7 \cdot X7 - 7104,98 \cdot X8 - 371,21 \cdot X9 - 63,0 \cdot X10 - 6757,98 \cdot X11 + 3415,81 \cdot X12 - 18,37 \cdot X13 + 716,99 \cdot X14,$$

$$F2 = -143859,0 + 3571,28 \cdot X1 - 563,80 \cdot X2 - 836,19 \cdot X3 - 22,82 \cdot X4 + 931,04 \cdot X5 + 21843,4 \cdot X6 + 1089,25 \cdot X7 - 7258,3 \cdot X8 - 381,62 \cdot X9 - 65,78 \cdot X10 - 6820,11 \cdot X11 + 3471,3 \cdot X12 - 18,39 \cdot X13 + 716,66 \cdot X14,$$

где  $X1$  — пропорциональность «н»;  $X2$  — контрастность «н»;  $X3$  — засечки «н»;  $X4$  — площадь «н»;  $X5$  — периметр «н»;  $X6$  — отношение кегля к высоте знака «н»;  $X7$  — отношение основного штриха к внутрибуквенному просвету «н»;  $X8$  — пропорциональность «а»;  $X9$  — контрастность «а»;  $X10$  — отношение максимальной ширины штриха к минимальному для знака «а»;  $X11$  — отношение кегля к высоте знака «а»;  $X12$  — отношение основного штриха к внутрибуквенному просвету «а»;  $X13$  — площадь «а»;  $X14$  — периметр «а».

Далее в качестве исходных условия было задействовано деление шрифтов на группы по методу скорости чтения [2].

В группу с высокой удобочитаемостью по методу скорости чтения включены 13 шрифтов: NewBaskervilleCTT, NewtonCTT, PragmaticaCTT, RaleighCTT BT, Times New Roman, CooperCTT Lt BT, FranklinGothicBookCTT, Gothic725CTT Bd BT, Humanist531CTT BT, SwiftCTT, TextBookCTT, OfficinaSansMediumCTT, OriginalGaramondSCCTT BT.

Группу с низкой удобочитаемостью по методу скорости чтения образовали 13 шрифтов: AdonisCTT, AvantGardeGothicCTT, BookmanCTT, BellGothicCTT BT, CharterCTT, ErasLightCTT, FuturaFuturisCTT, GeoSlb712CTT Md BT, KabelCTT-Medium, KisSCCTT BT, OCRF-RegularCTT, OctavaCTT, ZapfElliptical711CTT BT.

В данном случае была получена также одна дискриминантная функция с канонической корреляцией (0,91) и уровнем значимости функции по критериям Лямбда Уилкса (0,17) и хи-квадрат (30,19). Данные статистические показатели приблизительно равны полученным выше, и также являются удовлетворительными.

Также были рассчитаны коэффициенты, определяющие дискриминацию шрифтов по методу скорости чтения (таблица 3).

Таблица 3. Коэффициенты дискриминантной функции при классификации шрифтов по методу скорости чтения

Параметр	Коэффициент	Параметр	Коэффициент
Пропорциональность «н»	4,84	Контрастность «а»	-3,31
Контрастность «н»	-0,43	Max/Min «А»	-1,39
Засечки «н»	-1,72	Кегль/Высота «а»	-21,18
Площадь «н»	-0,05	Штрих/Просвет «а»	27,74
Периметр «н»	2,04	Площадь «а»	-0,03
Кегль/Высота «н»	69,81	Периметр «а»	1,33
Штрих/Просвет «н»	-4,85	Константа	-671,26
Пропорциональность «а»	1,95		

В данном случае вклад в дискриминацию вносят параметры отношения кегля к высоте для знаков «н» и «а», а также отношение основного штриха к внутрибуквенному просвету знака «а», т. е. эти параметры влияют на удобочитаемость шрифта, определенную методом скорости чтения. При методе парных сравнений значимым параметром также являлась величина пропорциональности «н».

Достаточно высокая точность классификации достигнута и в данном случае (таблица 4). Из 26 шрифтов в группе с высокой удобочитаемостью по методу скорости чтения находятся 13 объектов, в группе шрифтов с низкой удобочитаемостью состоит 13 объектов. Точность распознавания объектов первой группы составила 100 % (распознаны все 13 шрифтов), второй — 92,3 % (из 13 шрифтов распознаны 12).

Таблица 4. Классификационная таблица (метод скорости чтения)

Исходная группа	Предсказанная группа		Итого
	1	2	
1	13 (100 %)	0	13
2	1 (7,7 %)	12 (92,3 %)	13

Для прогнозирования удобочитаемости (по методу скорости чтения) новых шрифтов по геометрическим параметрам были получены следующие классификационные функции:

$$F1 = -326239,0 + 5271,39 \cdot X1 - 988,30 \cdot X2 - 1445,85 \cdot X3 - 48,85 \cdot X4 + 1958,53 \cdot X5 + 59104,5 \cdot X6 - 1248,83 \cdot X7 - 2193,24 \cdot X8 - 1961,56 \cdot X9 + 764,63 \cdot X10 - 17066,5 \cdot X11 + 17560,1 \cdot X12 - 34,01 \cdot X13 + 1467,28 \cdot X14,$$

$$F2 = -323382,0 + 5250,77 \cdot X1 - 986,47 \cdot X2 - 1438,53 \cdot X3 - 48,63 \cdot X4 + 1949,85 \cdot X5 + 58807,4 \cdot X6 - 1228,2 \cdot X7 - 2201,53 \cdot X8 - 1947,5 \cdot X9 + 778,71 \cdot X10 - 16976,4 \cdot X11 + 17442,0 \cdot X12 - 33,89 \cdot X13 + 1461,63 \cdot X14,$$

где  $X_1$  — пропорциональность «н»;  $X_2$  — контрастность «н»;  $X_3$  — засечки «н»;  $X_4$  — площадь «н»;  $X_5$  — периметр «н»;  $X_6$  — отношение кегля к высоте знака «н»;  $X_7$  — отношение основного штриха к внутрибуквенному просвету «н»;  $X_8$  — пропорциональность «а»;  $X_9$  — контрастность «а»;  $X_{10}$  — отношение максимальной ширины штриха к минимальному для знака «а»;  $X_{11}$  — отношение кегля к высоте знака «а»;  $X_{12}$  — отношение основного штриха к внутрибуквенному просвету «а»;  $X_{13}$  — площадь «а»;  $X_{14}$  — периметр «а».

## Выводы

С помощью этих функций можно в дальнейшем классифицировать новые шрифты, измерив их геометрические параметры. Шрифт будет относиться к первому классу (высокая удобочитаемость), если  $F_1 > F_2$ . Шрифт будет относиться ко второму классу (низкая удобочитаемость), если  $F_2 > F_1$ .

Полученный результат является вполне удовлетворительным и позволяет предсказывать качество для вновь разрабатываемых или тестируемых шрифтов, являющихся средствами передачи информации как в электронном, так и в печатном виде.

## Список литературы

1. Дюк В., Эмануэль В. Информационные технологии в медико-биологических исследованиях / СПб.: Питер, 2003. 528 с.
2. Токарь О. В. Комплексная оценка удобочитаемости современных типографских шрифтов на допечатной стадии полиграфического производства: автореф. дис. ... канд. тех. наук : 05.02.13 / МГУП. М., 2006. 22 с.
3. Токарь О. В., Зильберглейт М. А. Методика ассоциативных правил для анализа качества шрифтов // Полиграфия: технология, оборудование, материалы : материалы VI заочной научно-практической конференции с международным участием. Омск. 2015. С. 86–90.
4. Токарь О. В., Зильберглейт М. А. Взаимосвязь геометрических параметров шрифта и объективной и субъективной удобочитаемости // Технологія і техніка друкарства: збірник наукових праць. Киев. 2015. № 2 (48). С. 47–54.
5. Токарь О. В. Влияние геометрических параметров и емкости шрифта на его удобочитаемость // Передача, обработка, восприятие текстовой и графической информации: сборник трудов конференции. Екатеринбург. 2015. С. 144–152.